

**Determination of Binding Sites of Cadherin Peptides on the EC1 domain of E-cadherin  
using NMR Spectroscopy**

By

Copyright 2014

Elinaz Farokhi

Submitted to the graduate degree program in Pharmaceutical Chemistry and the Graduate  
Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of  
Master of Arts.

---

Chairperson Teruna J. Siahaan

---

Thomas Tolbert

---

Zhuo (Michael) Wang

Date Defended: May 16<sup>th</sup>, 2014

The Thesis Committee for Elinaz Farokhi

certifies that this is the approved version of the following thesis:

Determination of Binding Sites of Cadherin Peptides on the EC1 domain of E-cadherin using  
NMR Spectroscopy

---

Chairperson Teruna J. Siahaan

Date approved: May16th, 2014

## Abstract

The objective of this work is to evaluate the binding mechanisms of synthetic cadherin peptides (ADTC7, ADTC9 and cHAVc3) to the EC1 domain of human E-cadherin. The binding sites of these peptides to the EC1 domain were determined by titrating the  $^{15}\text{N}$ -labeled EC1 domain with each peptide. The changes in the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC NMR spectra were observed upon peptide titration. The AutoDock Vina molecular modeling program was also used to predict the binding sites of each cadherin peptide on the EC1 domain. NMR data confirmed that there are three potential binding sites of ADTC7 on the EC1 domain; these binding sites are around residues I53/V48, D103, and G115. The two potential binding sites of ADTC9 peptide are around I4 and I53/V48 residues of EC1. The NMR data showed that ADTC7 peptide has higher binding affinity to the EC1 domain than ADTC9 peptide. For cHAVc3, it binds selectively to at the D103 residue on EC1. AutoDock Vina studies confirm some of the binding sites found by NMR studies. In the future, a more detail study will be conducted to evaluate the effect of peptide binding on the dynamics properties of the EC1 domain.

# 1. Introduction

Cadherins are a large family of calcium-dependent adhesion molecules found in different tissues and in biological barriers such as the blood brain barrier (BBB) and intestinal mucosa [1]. They comprise at least five major subfamilies, which include type I classical cadherins, closely related type II cadherins, desmosomal cadherins (desmocollins and desmogleins), protocadherins, and a variety of cadherin-related molecules [2, 3]. E-cadherin is a type I cadherin found in epithelial tissues [2, 4, 5]. VE-cadherin is also found in endothelial cells and both E- and VE-cadherins are found in the microvessel endothelial cells of the blood-brain barrier [6]. E-cadherin-mediated reaggregation of bovine brain microvessel endothelial cells (BBMEC) can be inhibited by anti-E-cadherin antibody [6], suggesting the presence of E-cadherin in BBMEC. It has been indicated that endothelial cells in the brain are the tightest in the body while endothelial cells in bone marrow are more loose [7]. E-cadherin consists of a cytoplasmic tail, which contains the juxta-membrane domain (JMD) and catenin-binding domain (CBD), a trans-membrane region (T), and an extracellular region consisting of five repeated domains called EC1–5 [8]. EC1 and EC2 are important for cell-cell adhesion [9-11]. E- and VE-cadherins are responsible for cell-cell adhesion in the intercellular junctions of the BBB [12, 13]. The intercellular junctions contain three major parts: 1) *zonula occludens* (tight junctions), 2) *zonula adherens* (adheren junctions), and 3) *macula adherens* (desmosomes). E- and VE-cadherins are located in the adherens junctions [14]. Removal of calcium from the extracellular region of E- and VE-cadherins can loosen the intercellular junction of the cells [15, 16].

One of the major problems in treating brain disorders such as tumors and Alzheimer's and Parkinson's diseases is the difficulty in delivering drugs to the brain; this is due to the existence of the BBB [17]. Drugs cannot readily penetrate through the BBB via transcellular or

paracellular pathways. Drug penetration through the transcellular pathway occurs via partition of the drug into the cell membranes; thus, the membrane partition is influenced by the physicochemical properties of the drug. In addition, the presence of efflux pumps (e.g., P-glycoprotein (Pgp)) prevents the drug from crossing the transcellular pathway of the BBB. Alternatively, the drug can penetrate using the paracellular pathway; however, drug penetration through this pathway is limited by the presence of the tight junctions of the BBB [18, 19]. These normally restrict the crossing of hydrophilic molecules with a molecular weight higher than 180 Da [20].

The binding between the complex of natural killer cell receptor KLRG1 and the EC1 domain of E-cadherin has been determined using X-ray crystallography and residue mutation experiments [21]. Solution NMR and X-ray crystallography of ECAD12 were studied to evaluate the binding activity of the N-terminus of E-cadherin. At low protein concentration, intramolecular binding of tryptophan-2 (W2) side-chain into the W2 pocket on the same molecule was detected via chemical shift changes. However, at high protein concentration, the X-ray structure indicates that two molecules of ECAD12 form a dimer using intermolecular exchange of W2 side-chains that are docked to the W2-binding pockets of the opposite molecules [22].

Interactions of cadherins with proteins in the cytoplasmic domain are important for the structural stability and cell-cell adhesion properties of cadherins on the cell surface. For example, the complex of P120 catenin and cadherin contributes to cadherin surface stability and its role in cell-cell adhesion. The crystal structure of P120 with the JMD core region of E-cadherin has been shown using NMR, and this study confirms that P120 regulates the stability of cadherin-mediated cell-cell adhesion by binding to the majority of the JMD region [23].

Our group has found that cadherin peptides can modulate the intercellular junctions of the BBB to improve drug delivery into the brain. His-Ala-Val (HAV) and Ala-Asp-Thr (ADT) peptides derived from the EC1 domain can modify the porosity of the intercellular junctions of Madin-Darby canine kidney (MDCK) cell monolayers [24, 25]. These peptides lower the transepithelial electrical resistance (TEER) and enhance permeation of a paracellular marker molecule (i.e.,  $^{14}\text{C}$ -mannitol) and an anticancer drug ( $^3\text{H}$ -daunomycin) through the BBB in an *in situ* perfused rat brain [26]. Recently, HAV- and ADT-peptides have been shown to improve the *in vivo* brain delivery of marker molecules (i.e., R800 near-IR dye and gadopentetic acid) into the brain [27]. Binding studies of EC5 with HAV and BLG4 peptides were conducted, and the structural characteristics of EC5 were determined via NMR, CD, FTIR, and docking molecular modeling [19].

<b>Table 1. Peptides used in the binding studies to the EC1 domain of E-cadherin</b>		
<b>Peptide Name</b>	<b>Peptide Sequence</b>	<b>Molecular Weight (Da)</b>
<b>ADTC5</b>	Cyclo(1,7)Ac-CDTPPVC-NH <sub>2</sub>	744
<b>ADTC7</b>	Cyclo(1,7)Ac-CDTPPEC-NH <sub>2</sub>	802.89
<b>ADTC9</b>	Cyclo(1,7)Ac-CDTPPYC-NH <sub>2</sub>	836.95
<b>HAVc3</b>	Cyclo(1,6)Ac-CSHAVC-NH <sub>2</sub>	657
<b>HAV6</b>	Ac-SHAVSS-NH <sub>2</sub>	627.66

The objective of this work was to study the mechanism(s) through which peptides are able to modulate the paracellular pathway. The hypothesis is that the peptide activity is due to its

binding to the extracellular domain of E-cadherin. To confirm this, the potential binding sites of cyclic peptides on the EC1 domain of E-cadherin were investigated using 2D-NMR and molecular AutoDock program. Previously, the physical stability of the EC1 domain under different conditions (e.g., pH, temperature, ionic strengths, and buffer) was studied using CD, UV, and fluorescence spectroscopies [28]. The physical stability of EC1 can be improved by alkylation of the only cysteine residue on the EC1 domain [28]. The alkylation prevents the formation of a covalent dimer via an intermolecular disulfide bond of the cysteine residue on EC1. Recently, the NMR assignments of the amino acids within the human EC1 domain of E-cadherin have been completed using multidimensional  $^1\text{H}$ ,  $^{15}\text{N}$ , and  $^{13}\text{C}$  NMR spectra [29]. Using these assignments, the binding sites of several E-cadherin cyclic peptides such as ADTC7, ADTC9, and HAVc3 (Table 1) were determined. In this case, the  $^{15}\text{N}$ -labeled EC1 was titrated with cyclic peptides at different concentrations, the chemical shift changes at different residues were observed, and the magnitudes of the changes were determined as the overall chemical shift changes ( $\Delta F$ ). The chemical shift changes ( $\Delta F$ ) at a certain residue were plotted against peptide concentrations to calculate the dissociation constants ( $K_d$ ) of the peptide on a site on the EC1 domain. Molecular docking experiments were used to build working models for the most probable binding site(s) of each peptide on the EC1 domain of E-cadherin.

## ***2. Materials and Methods***

### *2.1. Transformation of pASK-IBA6-EC1 DNA into BL21 cells*

BL21 cells and plasmid cDNA vector were removed from  $-80^\circ\text{C}$  and  $-20^\circ\text{C}$  freezers, respectively, and thawed to  $0^\circ\text{C}$  on ice. The cDNA vector (1.0  $\mu\text{L}$ ) was added into 50  $\mu\text{L}$  cells,

and the mixture was incubated for 30 min on ice. Then, the cells were placed in a 42°C water bath for 30 sec followed by immediate transfer into ice for 3 min; this process was repeated several times to induce heat shock for permeation of the cDNA vector. Under sterile conditions, 200 µL of SOC medium (1.55 g yeast, 0.25 mL of 1 M KCl, 0.5 mL of 1 M MgCl<sub>2</sub>, 0.5 mL of 1 M MgSO<sub>4</sub>, 1 mL of 1 M glucose in distilled H<sub>2</sub>O) was added to each mini tube containing cells and plasmid cDNA vector, and the cells were agitated at 250 rpm for 1 h at 37°C. The culture tubes were removed from the shaker, and 50 µL and 150 µL of cells, respectively, were incubated into two agar plates. The agar-containing cells were placed in a standing incubating chamber to grow for 12–16 h. Next, the plates were taken from the incubation chamber, and each colony in the plate was picked up with a pipette tip and transferred into a 50 mL tube containing 10 mL of LB medium and 10 µL of ampicillin antibiotic (1000x). The cells in the tube were incubated in the shaker incubator overnight (12–16 h) at 250 rpm and 37°C. These cells were used as starter cells to produce the non-labeled and <sup>15</sup>N-labeled EC1 protein (see below).

## *2.2. Expression of <sup>15</sup>N-labeled EC1 protein using E. coli cells in M9 minimal medium*

To obtain the <sup>15</sup>N-labeled EC1 domain, the starter culture from the overnight incubation was added to a 1 L flask containing 250 mL of M9 minimal medium (34 g of Na<sub>2</sub>HPO<sub>4</sub>, 15 g of KH<sub>2</sub>PO<sub>4</sub>, and 2.5 g of NaCl supplemented with 5 g of <sup>15</sup>N-labeled ammonium chloride (<sup>15</sup>NH<sub>4</sub>Cl)). The cell growth was monitored at OD<sub>550</sub>, and 24 µL of anhydrotetracycline (2 mg/mL; Promega Inc., Madison, WI) was added to induce protein expression after the culture reached an OD of 0.5–0.6. This process usually takes about 3–5 h. The cells were incubated at 30°C for 6 h and then harvested by centrifugation at 4500 × g to produce cell pellets, which were stored at –80°C.



### 2.3. Expression of non-labeled EC1 protein using *E. coli* cells in LB medium

To obtain the non-labeled EC1 domain, the starter culture was added to a 1 L flask containing 250 mL of LB medium and 250  $\mu$ L of ampicillin antibiotic (1000 $\times$ ). The cell growth was monitored at OD<sub>550</sub>, and 24  $\mu$ L of anhydrotetracycline (2mg/ml; Promega Inc.) was added to induce protein expression after the culture reached an OD of 0.5–0.6. This process usually takes about 2–3 h. Then, the cells were incubated at 30°C for 6 h, after which time they were harvested by centrifuge at 4500  $\times$  g. The resulting cell pellets were stored at –80°C until further use.

### 2.4. Purification of the EC1 protein

The <sup>15</sup>N-labeled and non-labeled EC1 were isolated from *E. coli* cells after lysing the cells by French press at 20,000 psi in 15 mL of binding buffer (buffer B) (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, and 0.02% NaN<sub>3</sub>) at pH 8.0. The cellular debris was removed by centrifugation at 21,000  $\times$  g at 4°C for 1 h. The supernatant containing soluble EC1 was added into an equal volume of buffer B. The mixture was centrifuged at 4,500 rpm for 20 min at 4°C in an Amicon Ultra concentrator (EMD Millipore, Billerica, MA) with a 10 kDa molecular weight cut-off. The concentrated protein solution was added into a StrepTactin II affinity column (size 5.0  $\times$  0.6 cm, GE Healthcare) equilibrated with buffer B at room temperature. The column was washed with buffer B to remove the unbound proteins. To isolate the desired EC1 protein, the column was eluted (flowrate 2 mL/min) with elution buffer (buffer E containing 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.02% NaN<sub>3</sub>, 2.5 mM desthiobiotin, and 1 mM DTT) at pH 8.0. The purity of each fraction of EC1 was determined using 4–12% Tris-Bis SDS-PAGE. The

fractions of eluted protein were pooled and concentrated to 1.5 mL using an Amicon Ultra concentrator. The quantity of protein was determined using a UV spectrometer with molar absorption coefficient of  $19,480 \text{ M}^{-1} \text{ cm}^{-1}$  as calculated from the composition of tryptophan and tyrosine residues at 280 nm. The final pure EC1 was stored in buffer E at 4°C.

## 2.6. NMR experiments

The 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were performed on a Bruker Avance 600 MHz NMR spectrometer equipped with a triple resonance probe. The spectra were acquired in 16 scans with 2048 points in  $^1\text{H}$  resonance and 128 increments in  $^{15}\text{N}$  resonance. An internal standard (4,4-dimethyl-4-silapentane-1-sulfuric acid or DSS) in  $\text{D}_2\text{O}$  was used as a reference for the proton chemical shifts. The NMR data were processed with NMRDraw to determine the peak positions and chemical shifts in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra.

## 2.7. Binding studies of $^{15}\text{N}$ -labeled EC1 protein with cyclic ADTC7, ADTC9 and HAVc3 peptides

The  $^{15}\text{N}$ -labeled EC1 protein (0.18 mM) was titrated with 0.4 mM stock of ADTC7 peptide at protein/peptide ratios of 1:0.2, 1:0.4, 1:0.8, 1:1, 1:2, 1:3, 1:4, 1:5, 1:8, 1:10; 1:15, and 1:21. For the ADTC9 peptide, 0.4 mM  $^{15}\text{N}$ -labeled EC1 protein was titrated with 2.0 mM stock of peptide at protein/peptide ratios of 1:0.2, 1:0.4, 1:0.8, 1:1, 1:2, 1:3, 1:4, 1:6, 1:8, 1:10; 1:15, and 1:20. Finally, 0.4 mM of the  $^{15}\text{N}$ -labeled EC1 protein was titrated with 2.0 mM of cyclic HAVc3 peptide at protein/peptide ratios of 1:1, 1:4, and 1:8. The overall shift ( $\Delta F$ ) after titration was the

combined changes between  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts; this shift was calculated using the following equation:  $\Delta F = [(\Delta\delta^1\text{H} \cdot 599.743)^2 + (\Delta\delta^{15}\text{N} \cdot 60.778)^2]$ .

### *2.8. Molecular docking experiments with Autodock Vina:*

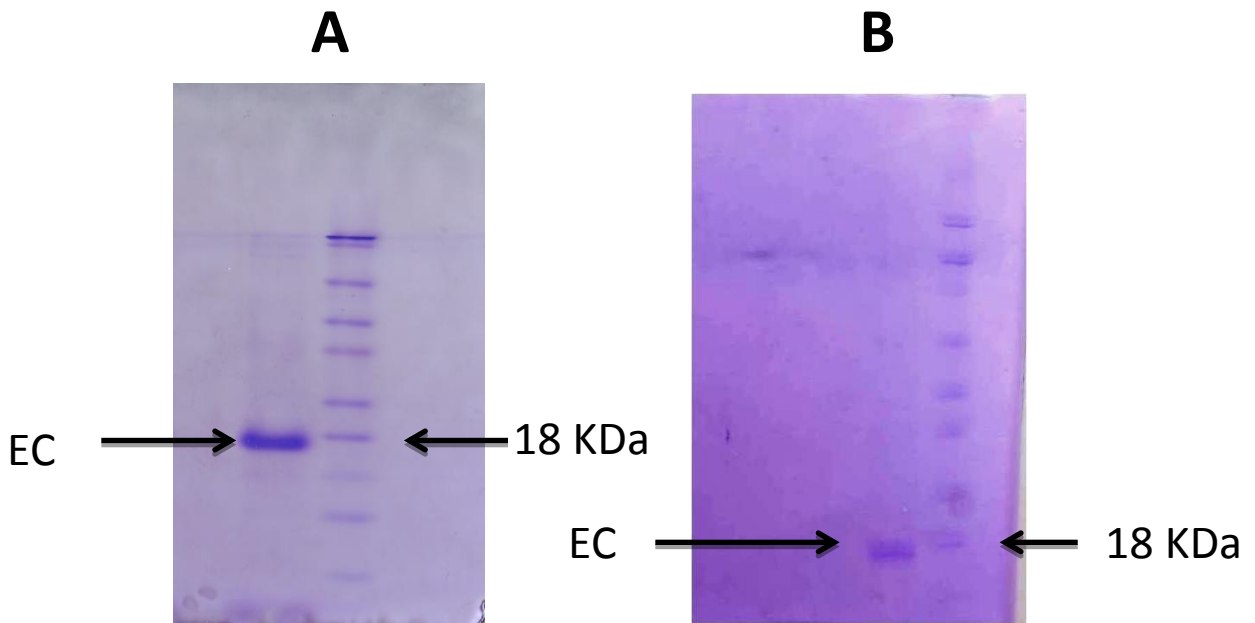
Molecular docking experiments were performed using Autodock Vina programs (Vina, The Scripps Institute) [30]. The EC1 domain structure used for the molecular docking experiments was derived from the X-ray structure of domain-1 and -2 of human E-cadherin acquired from the RCSB Protein Data Bank (entry code 2O72). While maintaining the same interface region, Discovery Studio Visualizer 2.5 program was used to remove water molecules and generate the EC1 domain model. The 3D structures of the peptides (i.e., ADTC5, ADTC7, ADTC9, and HAVc3) were generated using InsightII Program (Accelrys, Inc., San Diego, CA) and their conformations were optimized by energy minimization using CVFF91 force field. The final receptor and ligand coordinates were saved as protein data bank (pdb) files. After uploading these files into Autodock, AutodockTools were used to add partial charges and arrange the polar hydrogen in the protein. While maintaining the protein as a rigid structure, the ligands were set to have flexible torsion angles at all rotatable bonds. Using affinity energy (kcal/mol) and a root-mean-square deviation (RMSD) of 2.0, the interaction between peptide and EC1 was scored and ranked to generate potential binding sites.

## ***3. Results and Discussion***

### *3.1. Expression and purification of the EC1 domain*

Expression and purification of the non-labeled and  $^{15}\text{N}$ -labeled EC1 domain were accomplished without any problems. Figure 1 shows one band on the SDS-PAGE for purified non-labeled and  $^{15}\text{N}$ -labeled EC1 protein with a molecular weight of 18 KDa for EC1. The expressed EC1 domain has 150 amino acids of which (a) the first 13 amino acids at the N-terminus contain the MWSHPQFEKIEGR sequence for recognition by StrepTactin II affinity column, (2) the middle 110 amino acids are from the original sequence of the EC1 domain, and (c) the last 28 amino acids are from the interface region between the EC1 and EC2 domains (Figure 2).

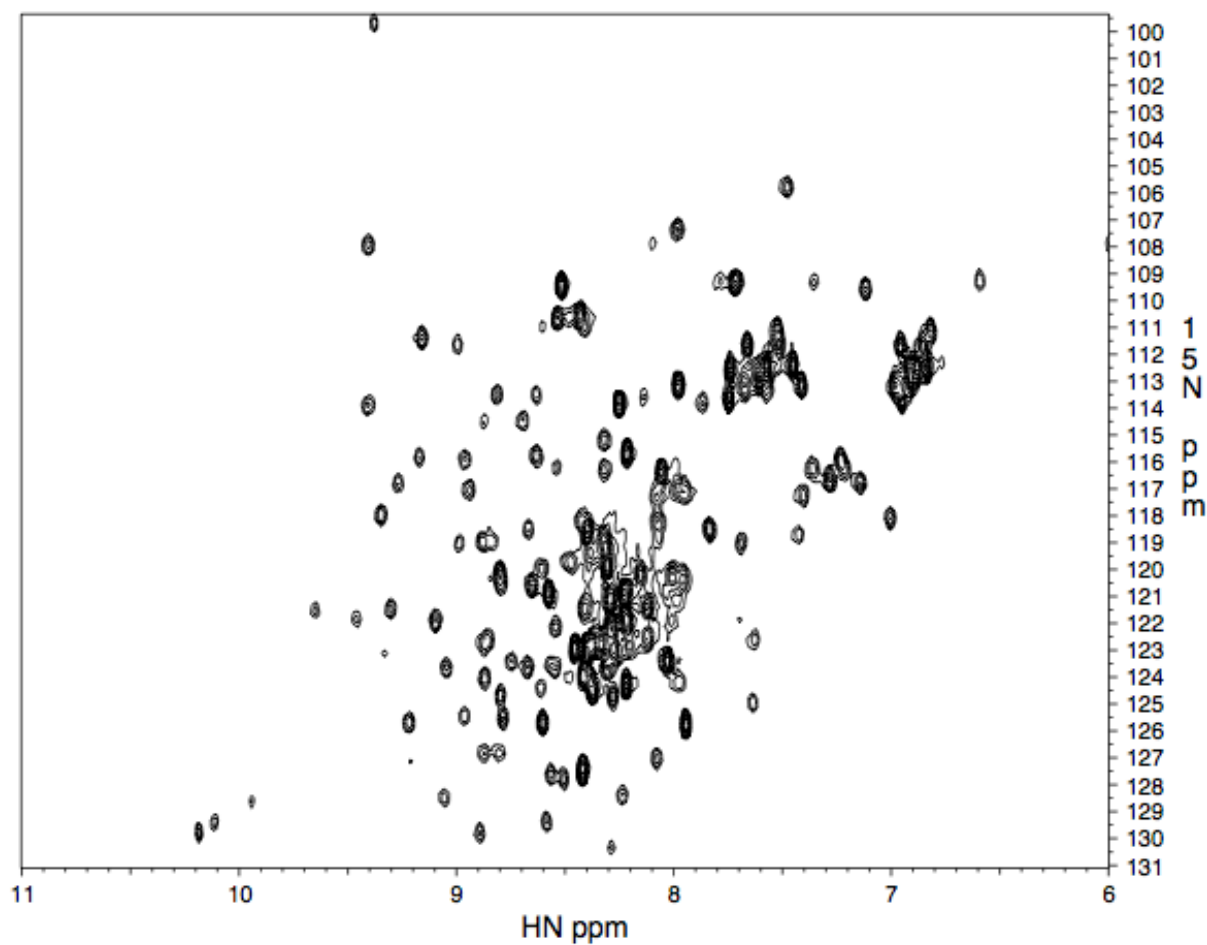
The  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC 2D-NMR spectra of the EC1 domain showed a good dispersion of crosspeaks from the amide groups, indicating the folded protein (Figure 2) [31]. The assignments of almost all of the residues have been completed; a few residues cannot be assigned due to overlapping resonances [29]. The residue assignments were used to evaluate the binding site(s) of peptide on the EC1 domain.



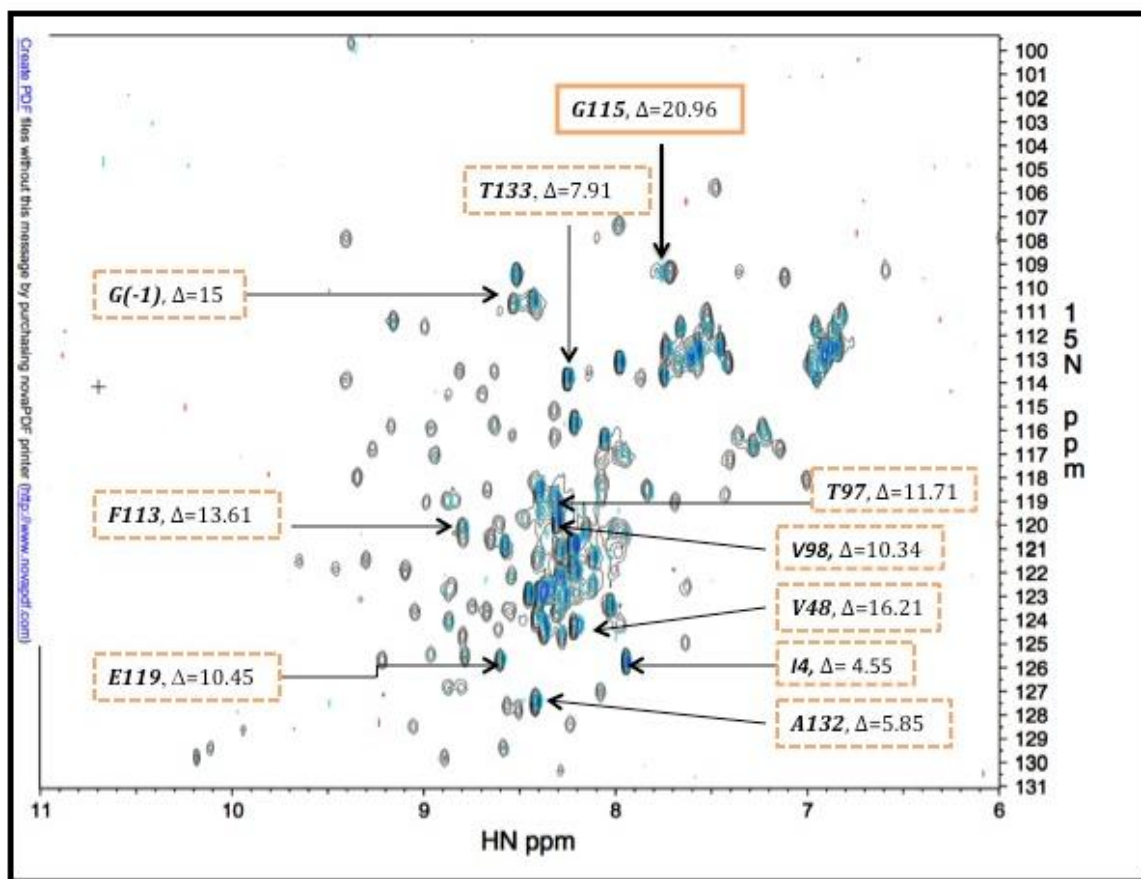
**Figure 1.** SDS-PAGE of the purified (A)  $^{15}\text{N}$ -labeled EC1 domain and (B) non-labeled EC1 domain of human E-cadherin after affinity column chromatography using a StrepTactin II column. The pure protein was run with the molecular weight marker molecules and the EC1 protein showed a molecular weight of 18 KDa.

**WSHPQFEKIEGRD** WV PP IS CP EN EK GP FP KN LV QK SN KD KE GK VF YS IT GQ  
 GADTPPVGVFII ER ET GW LK VT EP LD RE RI AT YT FS HA VS SN GN AV ED PM EI LL  
TV TD QN DN KP EFT Q EV FK GS VM EG AL PG TS VM EV TA TD AD DD

**Figure 2.** The amino acid sequence of the EC1 domain of human E-cadherin. The bold blue region is the tag sequence used for affinity purification of EC1 with a Strep-Tactin II affinity column. The black sequences are the 110 original amino acids for EC1. The green section is the sequence of the interface between the EC1 and EC2 domains. Most of the secondary structures in EC1 are beta sheets with one alpha helix. The sequences marked with red rectangles have beta-sheet secondary structure, and the one marked with a green cylinder shows the alpha-helix structure.



**Figure 3.**  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC 2D-NMR spectrum of the EC1 domain in 600 MHz NMR. The protein was dialyzed in 20 mM  $\text{KH}_2\text{PO}_4$  and 5 mM DTT buffer at pH 6.0, and 6%  $\text{D}_2\text{O}$  was added to 470  $\mu\text{L}$  of  $^{15}\text{N}$ -labeled EC1 sample.



**Figure 4.** Effect of titration of ADTC7 peptide on the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC 2D-NMR spectra of the EC1 domain observed using 600 MHz NMR. The protein (0.18 mM) was dissolved in 20 mM  $\text{KH}_2\text{PO}_4$  and 5 mM DTT buffer at pH 6.0. The spectrum of the protein alone is shown in black and the spectrum of the peptide-titrated protein (protein/peptide ratio = 1:15) is in blue. The data show the maximum changes in chemical shifts ( $\Delta F$ ) for residues that are affected by peptide addition.



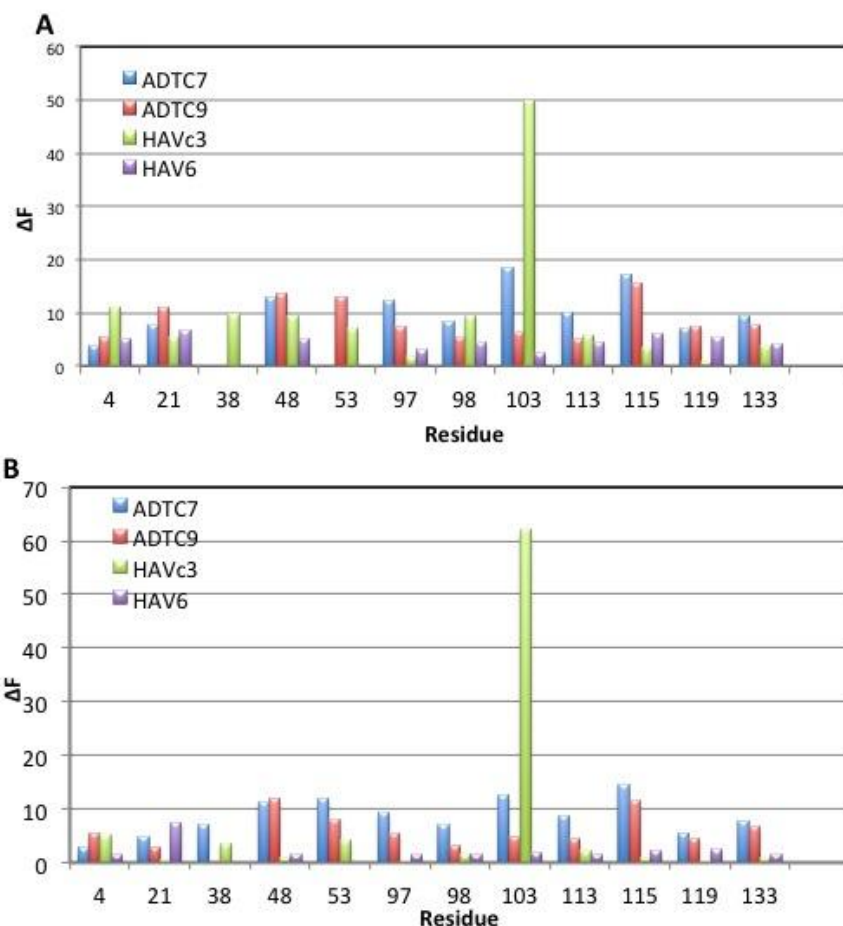
### *3.2. Studies of binding of ADTC7 to the EC1 domain*

#### *3.2.1 NMR binding studies*

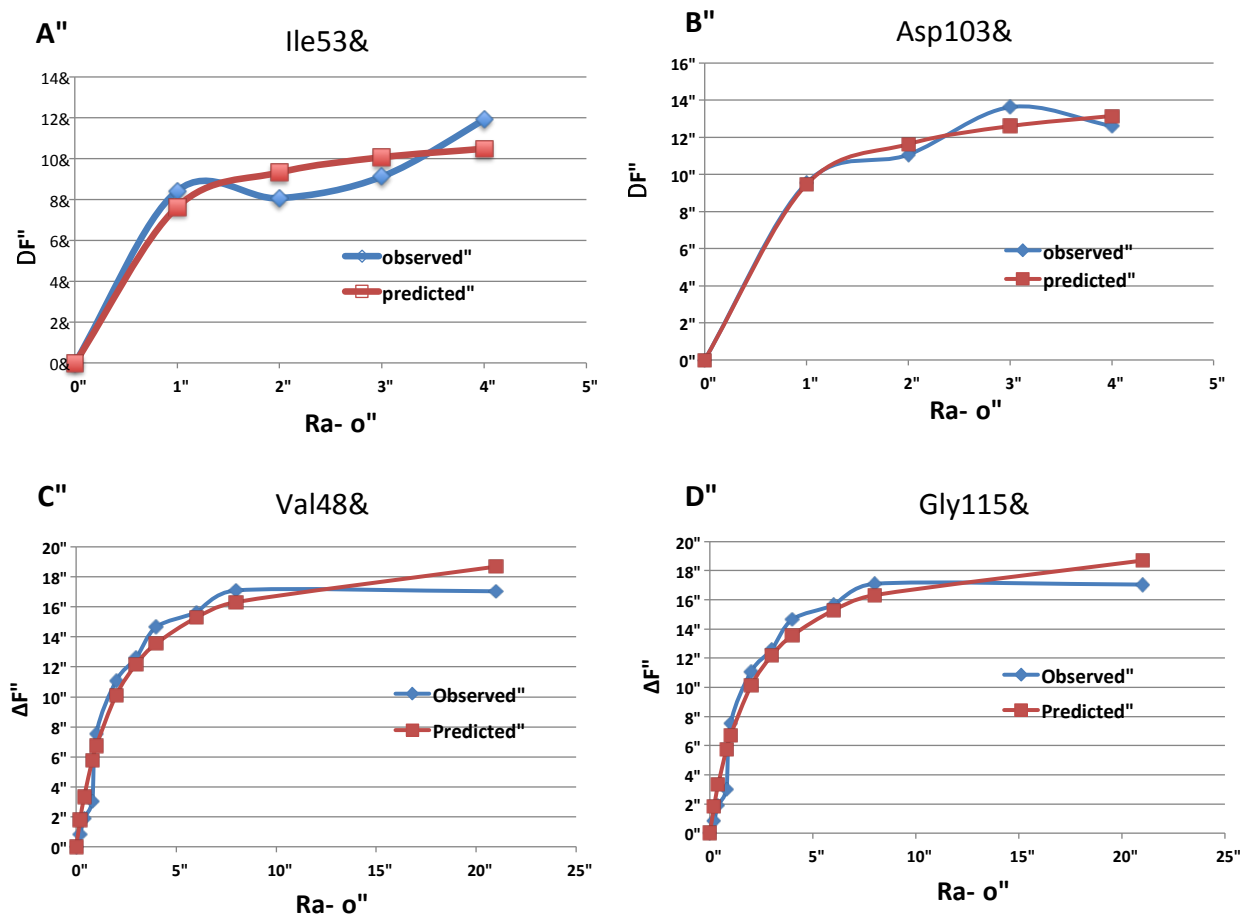
The binding site(s) of ADTC7 to the EC1 domain were evaluated using  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC 2D-NMR spectroscopy by titrating the protein with increasing concentrations of ADTC7 peptide with protein/peptide ratios between 1:0.2 and 1:20 (Figure 4). The crosspeaks shown in black are from the protein alone and the blue crosspeaks are from the protein titrated with the peptide. Upon titration with ADTC7, the NH crosspeaks were shifted with overall chemical shift changes in residues that have a high probability of binding to the peptide. These shifted crosspeaks were from the NH resonances of Ile4, Leu21, Ile38, Val48, Ile53, Thr97, Val98, Asp103, Phe113, Gly115, Glu119, Ala132, and Thr133 (Figure 5). These residues are from three different regions or sites of the EC1 domain and its C-terminal tail domain.

The magnitudes of changes at protein/peptide ratios of 1:4 and 1:8 are plotted in Figure 5. For EC1 titration with ADTC7, the NH crosspeaks of the Val48, Ile53, Asp103, and Gly115 residues in EC1 have large changes, suggesting that these residues are involved in the binding process between EC1 and ADTC7 peptide. Because these residues are located in different parts of EC1, an alternative explanation is that some residues that have shifted crosspeaks may be directly or indirectly involved in the binding process. In other words, the chemical shift changes may be due to conformational changes at the particular region of EC1. Therefore, there are several possible explanations for the changes in the chemical shifts at three different parts of the EC1 protein upon peptide addition. First, the peptide can bind to three different regions of the protein with similar or different binding affinities. Second, the peptide can bind to two regions of the protein that caused conformational changes in the third region of the protein, resulting in

chemical shift changes in the third region. Third, the peptide can bind to one site of the protein, resulting in conformational changes at two other regions of the protein and causing chemical shift changes in these remote sites or regions. At this time, the assumption is that any chemical shift change is due to binding between the peptide and the residue in different binding sites.



**Figure 5.** The overall chemical shift changes ( $\Delta F$ ) for several residues in the EC1 domain after addition of peptides at (A) protein/peptide ratio = 1:8 and (B) protein/peptide ratio = 1:4. For ADTC7 and ADTC9, the biggest shifts were on residues Gly115 and Val48. For HAVc3, the biggest change was on Asp103 and the second biggest change was on the Ile4 residue on the EC1 domain protein.



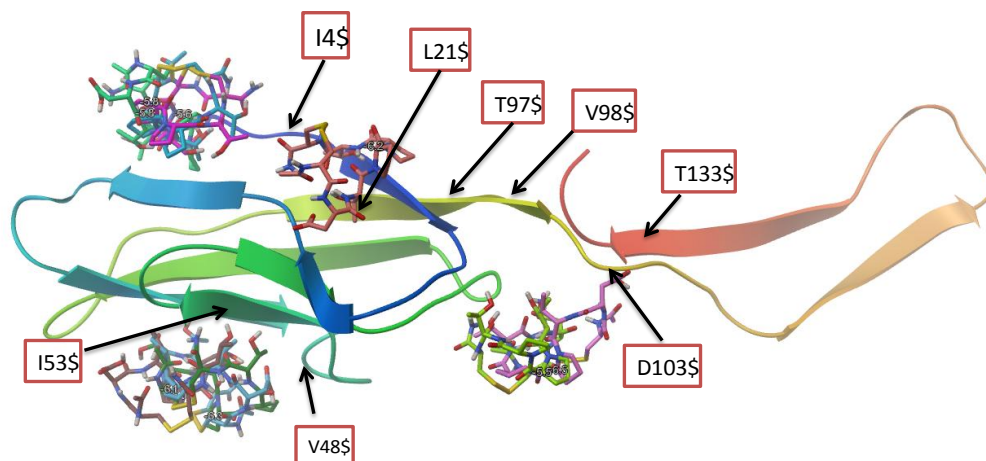
**Figure 6.** The effect of titration of different concentrations of ADTC7 peptide on the overall chemical shifts of (A) Ile53, (B) Asp103, (C) Val48, and (D) Gly115 residues on the EC1 domain. The peptide titration curves show binding saturation. The estimated  $K_d$  for ADTC7 binding to EC1 are (A)  $64.35 \mu M$  ( $R^2 = 0.97$ ) and (B)  $66.68 \mu M$  ( $R^2 = 0.98$ ) for Asp103, (C)  $137 \mu M$  ( $R^2 = 0.99$ ) for Val48, and (D)  $162 \mu M$  ( $R^2 = 0.96$ ) for Gly115.

To evaluate the tightness of binding of ADTC7 at different sites of EC1, the titration data were plotted to estimate the peptide dissociation constants at different sites of EC1. There are at least three potential binding constants of ADTC7 to EC1 domain. The first saturation curve was found at a protein/peptide ratio of 1:2, and the second saturation curve was at a protein peptide ratio of 1:7 on some residues of EC1. The first group of estimated dissociation constants ( $K_d$ ) for binding are around 64.35  $\mu\text{M}$  and 66.68  $\mu\text{M}$  as reflected in the Ile53 (Fig. 6A) and Asp103 (Fig. 6B) residues, respectively. The second group of dissociation constants are around 136–142  $\mu\text{M}$ , which are in the titration curves of Ile4, Val48, Thr97, and Val98 (Table 2). The third group of dissociation constants found were from Phe113 ( $K_d = 150 \mu\text{M}$ ), Thr133 ( $K_d = 150 \mu\text{M}$ ), and Gly115 ( $K_d = 162 \mu\text{M}$ ). These results suggest that the ADTC7 peptides could bind to three different regions of EC1 with different affinities.

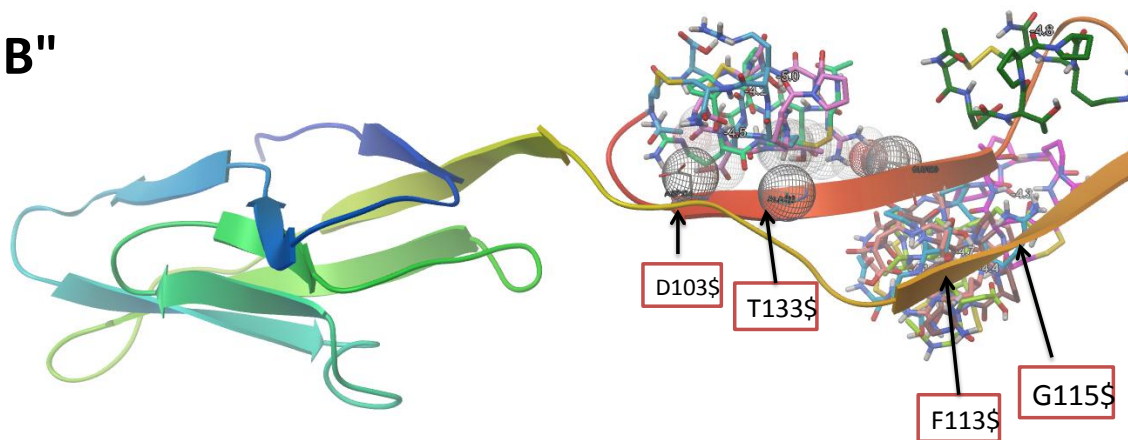
<b>Table 2. The Kd calculated from different residues in EC1 after titration with ADTC7 and ADTC9 peptides</b>		
	<b>Kd (<math>\mu</math>M)</b>	
<b>Residue</b>	<b>ADTC7</b>	<b>ADTC9</b>
<b>I4</b>	<b>142</b>	<b>403</b>
<b>L21</b>	<b>239.76(263.34)</b>	<b>181.81</b>
<b>I38</b>	<b>96.34(200.86)</b>	<b>-</b>
<b>V48</b>	<b>137</b>	<b>658</b>
<b>I53</b>	<b>64.35</b>	<b>218(464)</b>
<b>T97</b>	<b>136</b>	<b>1075</b>
<b>V98</b>	<b>137</b>	<b>965</b>
<b>D103</b>	<b>66.68(148)</b>	<b>450.21</b>
<b>F113</b>	<b>150</b>	<b>1031</b>
<b>G115</b>	<b>162</b>	<b>1023</b>
<b>E119</b>	<b>627</b>	<b>961</b>
<b>T133</b>	<b>150</b>	<b>814</b>

<b>Table 3. The chemical shift changes of residues in the EC1 domain upon addition of ADTC7 and ADTC9 peptides at a protein/peptide ratio of 1:8</b>			
	<b><math>\Delta F</math> (1:8)</b>		
<b>Residue</b>	<b>ADTC7</b>	<b>ADTC9</b>	<b>Cyclic HAVc3</b>
<b>I4</b>	<b>3.97</b>	<b>5.33</b>	<b>11.18</b>
<b>L21</b>	<b>7.78</b>	<b>11.02</b>	<b>5.35</b>
<b>I38</b>	<b>16.70</b>	<b>-</b>	<b>9.87</b>
<b>V48</b>	<b>13.03</b>	<b>13.59</b>	<b>9.36</b>
<b>I53</b>	<b>12.21</b>	<b>13.10</b>	<b>7.1</b>
<b>T97</b>	<b>12.46</b>	<b>7.33</b>	<b>1.84</b>
<b>V98</b>	<b>8.27</b>	<b>5.57</b>	<b>9.36</b>
<b>D103</b>	<b>18.57</b>	<b>6.50</b>	<b>49.95</b>
<b>F113</b>	<b>10.21</b>	<b>5.06</b>	<b>5.96</b>
<b>G115</b>	<b>17.1</b>	<b>15.56</b>	<b>3.63</b>
<b>E119</b>	<b>7.22</b>	<b>7.54</b>	<b>0.91</b>
<b>T133</b>	<b>9.44</b>	<b>7.87</b>	<b>3.8</b>

A''



B''



**Figure 7.** The results of docking experiments between ADTC7 and the EC1 domain and the tail region. (A) Four potential binding sites for ADTC7 on the EC1 domain were found. (B) Three potential binding sites of ADTC7 were found on the tail region of EC1.

### 3.2.2. Molecular docking of ADTC7 on EC1

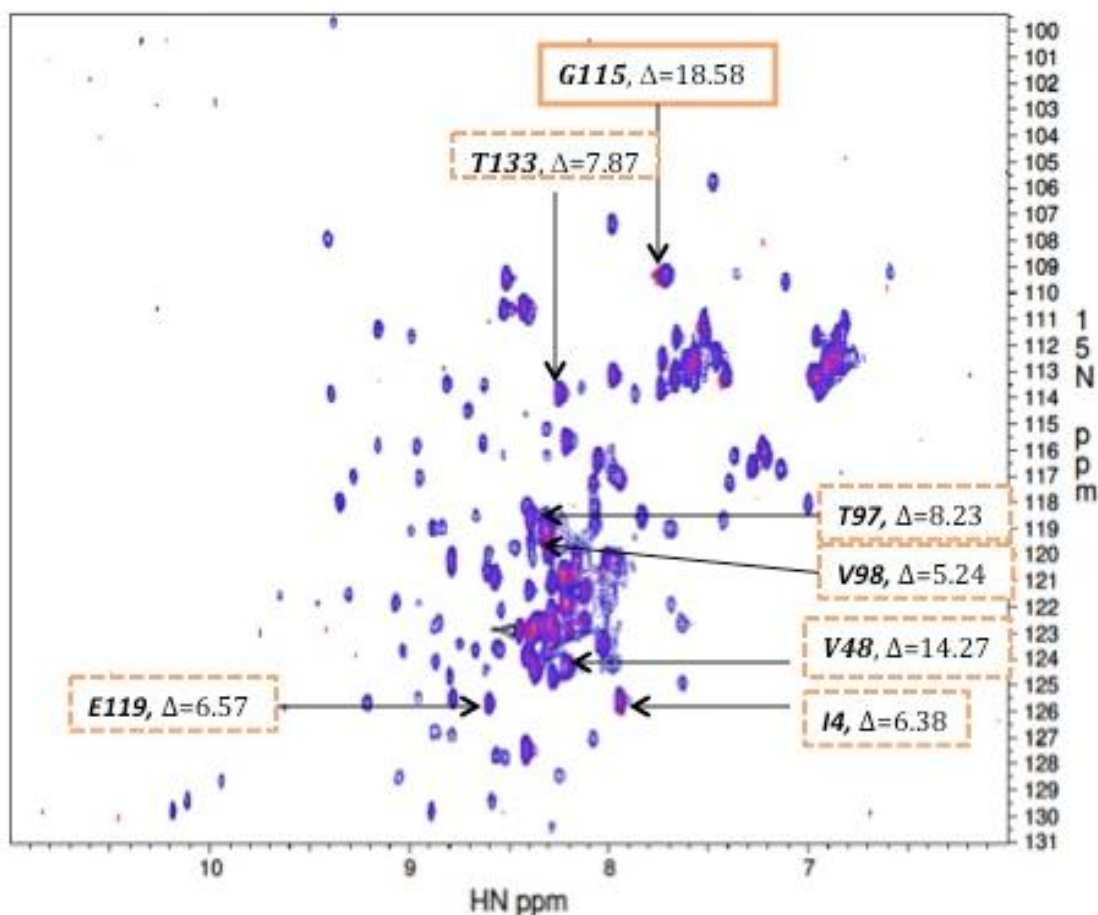
To evaluate potential binding sites of ADTC7 on the EC1 domain, molecular docking studies were carried out using the Autodock Vina program. In this study, the ADTC7 structure was built using the molecular modeling program Insight, and the peptide conformation was minimized using a Discover program. The coordinates of the peptides were converted to protein data bank (pdb) coordinates for molecular docking experiments. The coordinate for the EC1 domain was derived from the X-ray structure of EC1-2 domain of E-cadherin. Similar to the expressed EC1 domain, the model for EC1 domain contains the interface region between EC1 and EC2 domain, which is called the C-terminal tail region or tail region. The first and second docking experiments were done to search for peptide binding regions on the main EC1 structure and the C-terminal tail region, respectively.

Docking of ADTC7 to the main EC1 domain shows that the peptide binds to four different regions of EC1: the N-terminal, Leu21, Ile53, and Asp103 (Figure 7A). There are three potential binding sites for the tail region; the peptide can bind to Asp103/Thr133, Phe113/Gly115, and Glu119 (Figure 7B). The docking model of ADTC7 at Ile53 in the main EC1 was supported by the large chemical shift change of Ile53 (Figures 4 and 5) and the low K<sub>d</sub> (64  $\mu$ M, Figure 6A). The peptide docking at Ile53 also involves interactions with Ile38 and Val48, and the calculated K<sub>d</sub> using Ile38 and Val48 chemical shift changes are 96  $\mu$ M and 137  $\mu$ M, respectively. This result suggests that Ile38 and Val48 have weaker interactions with ADTC7 than does Ile53. There are several docking conformations in the Ile53 binding site with docking energies of -6.3, -6.1, and -6.0 kcal/mol.



The second favorable binding site of ADTC7 is at the Asp103 in the EC1 domain, which is supported by the larger chemical shift changes of the Asp103 residue during peptide titration (Figure 5, Table 3). The peptide docked at this site also binds to Thr133 residue. Binding to this site is predicted to also change the conformation at the Thr97 and Val 98 residues. Binding at the Asp103 site is also supported by the low  $K_d$  (66.68  $\mu\text{M}$ ) seen in the titration study. The involvements of Thr133, Thr97, and Val98 are supported by their  $K_d$  at 150  $\mu\text{M}$ , 136  $\mu\text{M}$ , and 137  $\mu\text{M}$ , respectively (Table 2). These values are higher than the  $K_d$  of Asp103, suggesting that they have a secondary role during binding at the Asp103 site.

The third possible binding model is peptide binding at the Phe113 and Gly115 residues of the tail region (Figure 7B). Upon titration with peptides, the NHs of Phe113 and Gly115 have large chemical shift changes. This binding site is less likely to be the primary binding site because the calculated  $K_d$  for binding to Phe113 and Gly115 are 150  $\mu\text{M}$  and 162  $\mu\text{M}$ . It is proposed that the tail region can be folded into the main EC1 domain upon binding to the peptide; thus, peptide binding to Phe113 and Gly115 can be the result of a secondary event in which a peptide binding to a primary site (i.e., Asp103) on the EC1 domain followed by folding of the tail region (Phe113 and Gly115) into the main region of EC1 around Asp103. Molecular dynamics simulations studies are being carried out to test this hypothesis.



**Figure 8.** <sup>1</sup>H,<sup>15</sup>N-HSQC 2D-NMR spectra of EC1 titrated with ADTC9 peptide observed using 600 MHz NMR. The protein was dissolved in 20 mM KH<sub>2</sub>PO<sub>4</sub> and 5 mM DTT buffer at pH 6.0 at 0.4 mM concentration. The spectrum of the protein is in purple and the spectrum of the peptide-titrated protein is pink; the protein/peptide ratio is 1:15. The data show the maximum change in chemical shifts (ΔF) for residues that are affected.

### *3.3. Binding studies of ADTC9 to the EC1 domain*

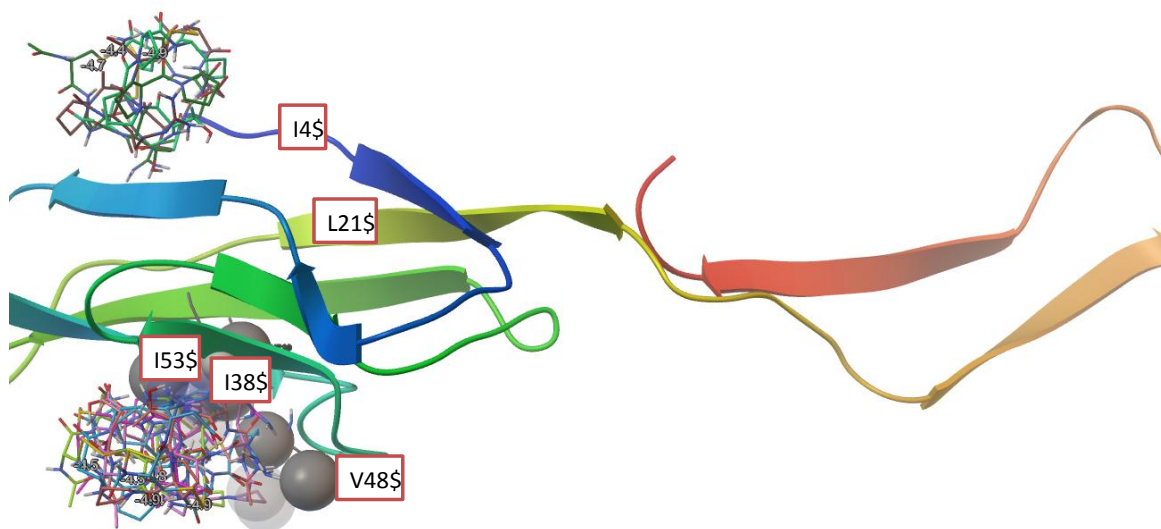
#### *3.3.1 NMR binding studies*

The EC1 domain was also titrated with ADTC9 peptide at protein/peptide ratios of 1:0.2 to 1:20, and the chemical shift changes ( $\Delta F$ ) were followed in various residues (Figures 5 and 8, Table 3). The greatest change was observed in the chemical shift of residues Leu21, Val48, Ile53, and Gly115. In contrast to ADTC7, ADTC9 did not cause a dramatic change in the chemical shift of Asp103; this is presumably due to mutation of the Glu6 residue in ADTC7 to the Tyr6 residue in ADTC9 (Table 3). The lowest  $K_d$  of binding of ADTC9 were found at Leu21 (181.81  $\mu\text{M}$ ) followed by Ile53 (218  $\mu\text{M}$ ). The titration curves for Leu21 and Ile53 have two saturation regions (i.e., at low and high peptide concentrations), suggesting the possibility of two binding sites. Unlike ADTC7, the calculated  $K_d$  at the Asp103 residue is high (450  $\mu\text{M}$ ), suggesting that the Asp103 residue is not a favorable binding site for ADTC9. Furthermore, the  $K_d$  for Thr133 was 814  $\mu\text{M}$ , which is very high compared to that of Thr133 (150  $\mu\text{M}$ ) for ADTC7 titration. Although the chemical shift changes were high for Gly115 when titrated with ADTC9, the  $K_d$  of Gly115 was very high (1023  $\mu\text{M}$ ), as were those of Phe113 (1031  $\mu\text{M}$ ) and Glu119 (961  $\mu\text{M}$ ). In summary, ADTC9 has two potential binding sites on the EC1 domain with dissociation constants of 181–218  $\mu\text{M}$  at Leu21 and 403–450  $\mu\text{M}$  at Ile53.

#### *3.3.2. Molecular docking of ADTC9 on EC1*

Molecular docking experiments showed that ADTC9 peptide has two binding sites on the main EC1 domain (Figure 9). One binding site is at the Ile53 residue and the other binding site is

at the N-terminal of EC1 around the Trp2 residue. Although the NMR data suggest that the Leu21 has the lowest  $K_d$  of binding, no ADTC9 was found docked at the Leu21 site on EC1. Thus, further studies are needed for evaluating the potential binding of ADTC9 at the Leu21 site on EC1. In contrast to ADTC7, ADTC9 does not bind to the tail region of EC1 nor does it interact with Gly115. Instead, the docking experiment suggests that ADTC9 interacts at another part of the tail region (i.e., Ala132, Asp134, Thr133 and Glu129).

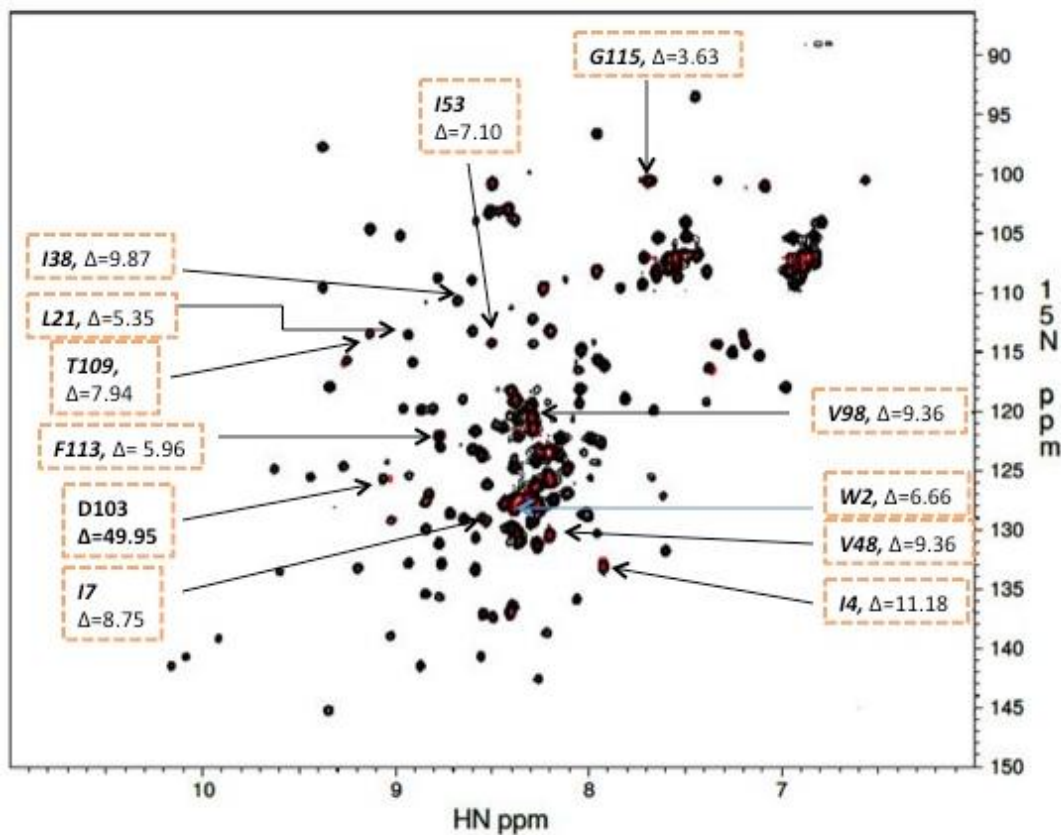


**Figure 9.** Molecular docking of ADTC9 to the EC1 domain. The two possible binding sites of ADTC9 on the main EC1 domain. The first binding site involves the Ile38, Ile53, and V48 residues, and the second involves the N-terminus of the EC1 domain. The second binding domain involved the Trp2 residue at the N-terminus.

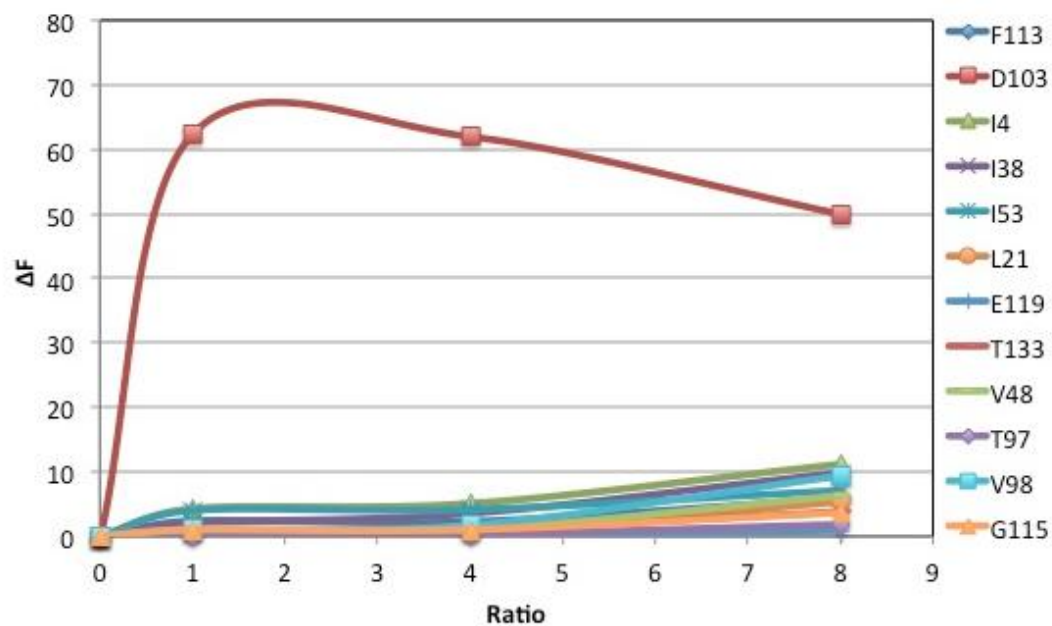
### *3.4. Binding studies of HAVc3 to the EC1 domain*

#### *3.4.1 NMR binding studies*

Similar to ADTC7 and ADTC9, the EC1 domain was titrated with different concentrations of HAVc3, and the chemical shift changes were observed in  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC 2D-NMR spectra (Figure 10). The largest chemical shift change ( $\Delta F$ ) was found at the Asp103 residue (Figures 5 and 10 and Table 3). The second biggest chemical shift changes were found at the Ile4 residue and the Ile38 residue. The chemical shift changes in Leu21 and Gly115 upon titration with HAVc3 were much lower than those produced by titration with ADTC7, suggesting that HAVc3 binds at different sites than ADTC7. The effects of peptide concentrations on the chemical shift changes are shown in Figure 11. Due to a limited number of titration concentrations, it was difficult to determine the  $K_d$  at affected residues on the EC1 domain. Because of the large change in the Asp103 chemical shift compared to those of other residues, the result suggests that HAVc3 selectively binds around the Asp103 residue and not to other residues or sites. This also suggests that HAVc3 has only one binding site at the EC1 domain.



**Figure 10.**  $^1\text{H}$ ,  $^{15}\text{N}$ - HSQC 2D-NMR spectra of EC1 titrated with cyclic HAVc3 peptide using 600 MHz NMR. The protein was dissolved in 20 mM  $\text{KH}_2\text{PO}_4$  and 5 mM DTT buffer at pH 6.0 at 0.4 mM concentration. The spectrum of the protein is in black and the spectrum of peptide-titrated protein is red; the protein/peptide ratio is 1:8. The data show the maximum change in chemical shifts ( $\Delta F$ ) for residues that are affected.



**Figure 11.** Binding curves for different residues of EC1 titrated to the HAVc3 peptide. The chemical shift of Asp103 shows large changes upon saturation at the 1:1 ratio, while other residues show low chemical shift changes.

## ***4. Conclusion***

The NMR studies indicate that ADTC7, ADTC9 and HAVc3 bind to the EC1 domain of E-cadherin. The ADTC7 has higher binding properties than ADTC9. The NMR data indicate that ADTC7 and ADTC9 have favorable binding sites at the Ile53 and Leu21 residues. ADTC7 has an additional binding site at the Asp103 residue while ADTC9 does not. The NMR data show that HAVc3 binds selectively to the Asp103 residue in EC1. The docking experiments suggested that ADTC7 and ADTC9 bind favorably to the Ile53 sites at the EC1 domain. In the future, more detailed NMR titration studies will be carried out on HAVc3 peptide at concentrations between protein/peptide ratios from 1:0.2 to 1:3. These studies are an attempt to increase the number of data points for K<sub>d</sub> determination. The docking experiments will also be carried out for the HAVc3 peptide. Finally, molecular dynamics simulation studies will be initiated to evaluate the effects of peptide binding on the conformational changes in the EC1 domain, especially the effect of peptide binding on the folding of the tail region of EC1.



## **5. References:**

- [1] B.M. Gumbiner, Regulation of cadherin-mediated adhesion in morphogenesis, *Nat Rev Mol Cell Biol*, 6 (2005) 622–634.
- [2] P. Hulpiau, F. van Roy, Molecular evolution of the cadherin superfamily, *Int J Biochem Cell Biol*, 41 (2009) 349–369.
- [3] C. Redies, K. Vanhalst, F. Roy, delta-Protocadherins: unique structures and functions, *Cell Mol Life Sci*, 62 (2005) 2840–2852.
- [4] D.G. Huntsman, C. Caldas, Assignment1 of the E-cadherin gene (CDH1) to chromosome 16q22.1 by radiation hybrid mapping, *Cytogenet Cell Genet*, 83 (1998) 82–83.
- [5] B.D. Angst, C. Marcozzi, A.I. Magee, The cadherin superfamily: diversity in form and function, *J Cell Sci*, 114 (2001) 629–641.
- [6] D. Pal, K.L. Audus, T.J. Siahaan, Modulation of cellular adhesion in bovine brain microvessel endothelial cells by a decapeptide, *Brain Res*, 747 (1997) 103–113.
- [7] M. Tavassoli, M. Shaklai, Absence of tight junctions in endothelium of marrow sinuses: possible significance for marrow cell egress, *British journal of haematology*, 41 (1979) 303–307.
- [8] N. Ishiyama, S.H. Lee, S. Liu, G.Y. Li, M.J. Smith, L.F. Reichardt, M. Ikura, Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion, *Cell*, 141 (2010) 117–128.
- [9] A. Tomschy, C. Fauser, R. Landwehr, J. Engel, Homophilic adhesion of E-cadherin occurs by a co-operative two-step interaction of N-terminal domains, *EMBO J*, 15 (1996) 3507–3514.

- [10] T.J. Boggon, J. Murray, S. Chappuis-Flament, E. Wong, B.M. Gumbiner, L. Shapiro, C-cadherin ectodomain structure and implications for cell adhesion mechanisms, *Science*, 296 (2002) 1308–1313.
- [11] W. He, P. Cowin, D.L. Stokes, Untangling desmosomal knots with electron tomography, *Science*, 302 (2003) 109–113.
- [12] K. Zheng, M. Trivedi, T.J. Siahaan, Structure and function of the intercellular junctions: barrier of paracellular drug delivery, *Curr Pharm Des*, 12 (2006) 2813–2824.
- [13] P. Kiptoo, M.D. Laksitorini, T.J. Siahaan, Blood-Brain Peptides: Peptide Delivery, in: A. Kastin (Ed.) *Handbook of Biologically Active Peptides*, Academic Press, Boston, 2013, pp. 1702–1710.
- [14] B. Gumbiner, Structure, biochemistry, and assembly of epithelial tight junctions, *Am J Physiol*, 253 (1987) C749–758.
- [15] J.L. Bixby, R. Zhang, Purified N-cadherin is a potent substrate for the rapid induction of neurite outgrowth, *J Cell Biol*, 110 (1990) 1253–1260.
- [16] M. Takeichi, Cadherins: a molecular family important in selective cell-cell adhesion, *Annu Rev Biochem*, 59 (1990) 237–252.
- [17] K.L. Cepek, S.K. Shaw, C.M. Parker, G.J. Russell, J.S. Morrow, D.L. Rimm, M.B. Brenner, Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the alpha E beta 7 integrin, *Nature*, 372 (1994) 190–193.
- [18] K.L. Lutz, T.J. Siahaan, Molecular structure of the apical junction complex and its contribution to the paracellular barrier, *J Pharm Sci*, 86 (1997) 977–984.

- [19] K. Zheng, J.S. Laurence, K. Kuczera, G. Verkhivker, C.R. Middaugh, T.J. Siahaan, Characterization of multiple stable conformers of the EC5 domain of E-cadherin and the interaction of EC5 with E-cadherin peptides, *Chem Biol Drug Des*, 73 (2009) 584–598.
- [20] L.L. Mitic, J.M. Anderson, Molecular architecture of tight junctions, *Annu Rev Physiol*, 60 (1998) 121–142.
- [21] Y. Li, M. Hofmann, Q. Wang, L. Teng, L.K. Chlewicki, H. Pircher, R.A. Mariuzza, Structure of natural killer cell receptor KLRG1 bound to E-cadherin reveals basis for MHC-independent missing self recognition, *Immunity*, 31 (2009) 35–46.
- [22] D. Haussinger, T. Ahrens, T. Aberle, J. Engel, J. Stetefeld, S. Grzesiek, Proteolytic E-cadherin activation followed by solution NMR and X-ray crystallography, *EMBO J*, 23 (2004) 1699–1708.
- [23] A.B. Reynolds, Exposing p120 catenin's most intimate affair, *Cell*, 141 (2010) 20–22.
- [24] I.T. Makagiansar, M. Avery, Y. Hu, K.L. Audus, T.J. Siahaan, Improving the selectivity of HAV-peptides in modulating E-cadherin-E-cadherin interactions in the intercellular junction of MDCK cell monolayers, *Pharm Res*, 18 (2001) 446–453.
- [25] E. Sinaga, S.D. Jois, M. Avery, I.T. Makagiansar, U.S. Tambunan, K.L. Audus, T.J. Siahaan, Increasing paracellular porosity by E-cadherin peptides: discovery of bulge and groove regions in the EC1-domain of E-cadherin, *Pharm Res*, 19 (2002) 1170–1179.
- [26] P. Kiptoo, E. Sinaga, A.M. Calcagno, H. Zhao, N. Kobayashi, U.S. Tambunan, T.J. Siahaan, Enhancement of drug absorption through the blood-brain barrier and

- inhibition of intercellular tight junction resealing by E-cadherin peptides, *Molecular pharmaceutics*, 8 (2011) 239–249.
- [27] N.H. On, P. Kiptoo, T.J. Siahaan, D.W. Miller, Modulation of Blood-Brain Barrier Permeability in Mice Using Synthetic E-Cadherin Peptide, *Molecular pharmaceutics*, (2014).
- [28] M. Trivedi, J.S. Laurence, T.D. Williams, C.R. Middaugh, T.J. Siahaan, Improving the stability of the EC1 domain of E-cadherin by thiol alkylation of the cysteine residue, *Int J Pharm*, 431 (2012) 16–25.
- [29] V.D. Prasasty, M.E. Krause, U.S. Tambunan, A. Anbanandam, J.S. Laurence, T.J. Siahaan, H, C and N backbone assignment of the EC-1 domain of human E-cadherin, *Biomol NMR Assign*, (2014).
- [30] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J Comput Chem*, 31 (2010) 455–461.
- [31] I.T. Makagiansar, A. Ikesue, P.D. Nguyen, J.L. Urbauer, R.J. Urbauer, T.J. Siahaan, Localized production of human E-cadherin-derived first repeat in *Escherichia coli*, *Protein expression and purification*, 26 (2002) 449–454.